Optics and Lasers in Engineering 49 (2011) 521–525

Contents lists available at ScienceDirect

Optics and Lasers in Engineering

journal homepage: www.elsevier.com/locate/optlaseng

Optimization of the Hartmann–Shack microlens array

Otávio Gomes de Oliveira *, Davies William de Lima Monteiro

OptMAlab, Room 1212, DEE/UFMG, Universidade Federal de Minas Gerais, Av. Antonio Carlos, 6627 Pampulha, 31270-010 Belo Horizonte, Minas Gerais, Brazil

ARTICLE INFO

Article history:
Received 1 October 2010
Received in revised form
1 January 2011
Accepted 3 January 2011
Available online 15 January 2011

Keywords:
Adaptive optics
Microlens array
Lenslet array
Hartmann–Shack wavefront sensor

ABSTRACT

In this work we propose to optimize the microlens-array geometry for a Hartmann–Shack wavefront sensor. The optimization makes possible that regular microlens arrays with a larger number of microlenses are replaced by arrays with fewer microlenses located at optimal sampling positions, with no increase in the reconstruction error. The goal is to propose a straightforward and widely accessible numerical method to calculate an optimized microlens array for a known aberration statistics. The optimization comprises the minimization of the wavefront reconstruction error and/or the number of necessary microlenses in the array. We numerically generate, sample and reconstruct the wavefront, and use a genetic algorithm to discover the optimal array geometry. Within an ophthalmological context, as a case study, we demonstrate that an array with only 10 suitably located microlenses can be used to produce reconstruction errors as small as those of a 36-microlens regular array. The same optimization procedure can be employed for any application where the wavefront statistics is known.

1. Introduction

The Hartmann–Shack (H–S) wavefront sensor [1] is now deployed in many different fields, from astronomy to industrial inspection [2], where the quality of optical media or components can be measured by the distortions they impart on a wavefront transmitted or reflected by them. In ophthalmology, this sensor is a core component of major aberrometers, used to assess the visual quality of the eye and applied to academic research, real-time surgery monitoring and clinical diagnosis [3].

The microlens array is an important element in the H–S sensor, responsible for sampling the aberrated wavefront into light spots on the focal plane. The position of each light spot relates to the average tilt of the wavefront over the respective microlens. These spot-position coordinates are then used in the modal reconstruction [4] to approximate the wavefront topology with a combination of orthogonal basis functions, e.g. Zernike polynomials [5]. In the modal approach for wavefront reconstruction, the aberration $W(x,y)$ is approximated as a linear combination of a finite number of Zernike polynomials, as in Eq. (1)

$$W(x,y) = \sum_{i=0}^{M} c_i Z_i(x,y), \tag{1}$$

where c_i are the Zernike coefficients, $Z_i(x,y)$ the orthogonal Zernike functions and M is the number of Zernike terms used in this truncated representation. An arbitrary incoming wavefront sampled by N microlenses yields N spots on the focal plane displaced from their respective reference positions. Each spot displacement x and y are related to the corresponding x and y slopes averaged over each sampling microlens. These slopes S_i can be arranged as a $2N \times 1$ vector that relates to a vector of M Zernike coefficients c_i through a $2N \times M$ reconstruction matrix **B**, as in Eq. (2):

$$\mathbf{S} = \mathbf{Bc}. \tag{2}$$

Matrix **B** contains the averaged values of first derivatives of the Zernike functions over each sampling region.

The least-squares method can be used to estimate the Zernike coefficients, through the computation of the Moore–Penrose pseudoinverse of **B** (Eq. (3))

$$\mathbf{A} = (\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T, \tag{3}$$

where **A** will be henceforward referred to as the estimation matrix. Therefore, the estimation of **c** reduces to solving the set of equations represented by Eq. (4)

$$\mathbf{c} = (\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T\mathbf{S} \tag{4}$$

where $\mathbf{B}^T$ is the transpose of **B**.

The wavefront sampling is influenced by the microlens distribution pattern, lens contour and size, number of microlenses and fill factor. Adopted grids typically consist in either rectangular or hexagonal configurations. Soloviev and Vdovin [6] have discussed the influence of the geometry of a microlens array on the wavefront reconstruction error. They proposed a generic mathematical model to describe the modal wavefront reconstruction,

* Corresponding author. Tel.: +55 31 3409 3497; fax: +55 31 3409 4810.
E-mail address: otaviogomes@cpdee.ufmg.br (O.G. de Oliveira).

0143-8166/© 2011 Elsevier Ltd. Open access under the Elsevier OA license.
doi:10.1016/j.optlaseng.2011.01.002

independent of both the basis functions and the wavefront statistics. Their model indicates the dependency of the reconstruction error on the array geometry. To validate their assumptions, they evaluated different array geometries with Zernike coefficients from the atmospheric-turbulence statistics [7]. Arrays with randomly distributed microlenses generated lower reconstruction errors than regular grids, especially when the number of Zernike terms increased beyond 40, for which there was a catastrophic growth of the sampling error.

Based on these results we further investigate a methodology to identify an optimized microlens array for a known scenario to minimize either the reconstruction error or the number of microlenses, as compared to regular arrays. Here we apply the methodology to the context of ophthalmology, as a case study. The importance of specifying an optimal array for this application is expressed in the words of Llorente et al. [3]: "The determination of a sampling pattern with the minimum sampling density that provides accurate results is of practical importance for sequential aberrometers, since it would decrease measurement time, and of general interest to better understand the trade-offs between aberrometers. It is also useful to determine whether there are sampling patterns that are better adapted to typical ocular aberrations, or particular sampling patterns optimized for measurement under specific conditions." Note that, besides determining an optimal array for a specified number of microlenses, it is also desirable to determine arrays with a smaller number of microlenses, preserving the reconstruction error magnitude, which is exploited in this work.

2. Methodology

The search for the optimal microlens array is carried out numerically. Given an array, an algorithm generates input wavefronts based on statistics for typical ocular aberrations and an H–S model performs its sampling and modal reconstruction. The outcome is the root-mean-square (rms) error between the reconstructed wavefront and the input one. This procedure is coupled to an optimization method where the H–S model represents the objective function. Since our emphasis lies on the reconstruction error due to the optical sampling at the microlens array, errors related to photodetection, spot-coordinate detection and electronic sampling are not taken into account. In a real situation, however, these sources of errors cannot be neglected, since their magnitude can be greater than the sampling error. This emphasizes again the convenience of determining arrays with fewer microlenses. They contribute to a reduction in the overall error associated with the aforementioned noise sources, while preserving the sampling error as low as possible. Moreover, the array with fewer elements affords a design with larger microlenses. In this case, more light impinges on each microlens, therefore increasing the signal-to-noise ratio. A smaller number of sampled light spots yields fewer coordinate data which contributes, particularly, to smaller error propagation, regardless the reconstructor used. The method searches the microlens array geometry that minimizes the reconstruction error, as illustrated in Fig. 1.

The H–S model (C++ code), extended a source code initially developed at the EI Lab/TU-Delft/The Netherlands by G. Vdovin, D.W. de Lima Monteiro and S. Sakarya [4], in which the reconstruction is based on the least-square approximation [8]. A model of the microlenses was included in the code to simulate the sampling step based on the geometrical description of a microlens, as shown in Fig. 2. In this model the measured slope corresponds to the average slope calculated in 45 points distributed over the microlens area and is represented in the figure by the angle θ.

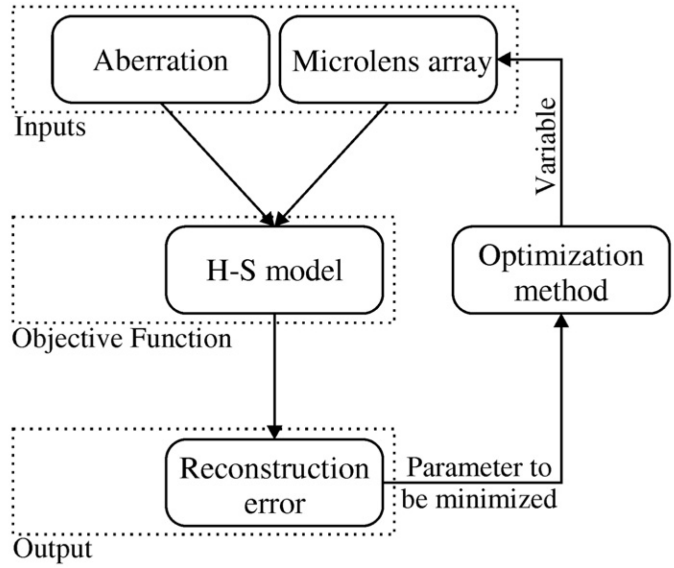

Fig. 1. Representation of the microlens array optimization process.

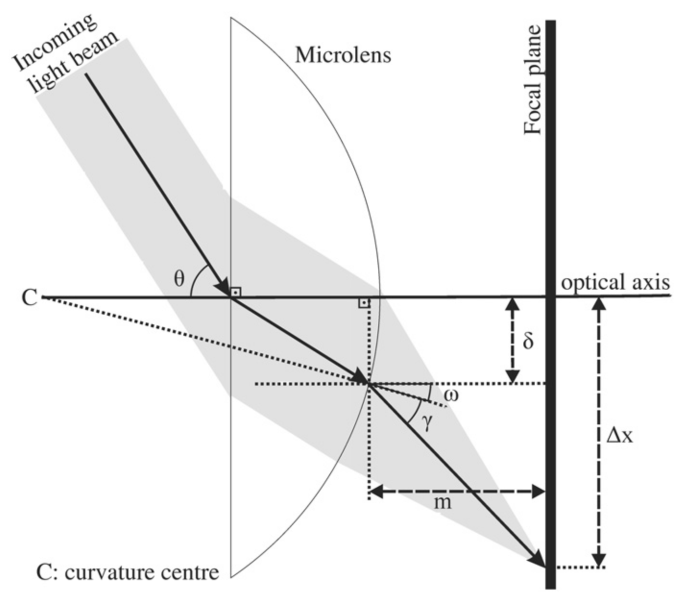

Fig. 2. Geometrical optics model for a microlens.

The light-spot deviation Δx is calculated through Eq. (5)

$$\Delta x = m\tan(\gamma+\omega)+\delta, \tag{5}$$

where the variables are identified in Fig. 2.

The wavefront reconstruction error from the H–S model results from the combination of two non-correlated sources: sampling and least-square modal reconstruction. It is not possible to separate the sampling error from numerical round-off errors introduced at the reconstruction routine. Therefore, henceforth the combination of these error components will be referred to as the reconstruction error.

The statistics for ocular aberrations was obtained from the work in which Porter et al. [9] measured the wavefront aberrations of both eyes of 109 human subjects, describing it in terms of 18 Zernike coefficients, with piston and tilts removed. They listed the typical mean values and spread of the Zernike coefficients for a mean pupil diameter of 5.7 mm. This data was used as a basis for the wavefront-generator algorithm (WG). Therefore, in this work, the simulations use the first 20 Zernike modes, once that is the maximum number of modes with available statistics in the work of Porter et al. [9]. Moreover, de Lima Monteiro [10] states that optical aberrations in the human eye are significant only to the first 14th Zernike modes, including tip and tilt. However, tip, tilt and defocus were not considered, so that the microlens array could be optimized to provide a more precise sampling of higher order Zernike terms. More reconstruction modes could be used, if they were known by the statistics. It is relevant to comment here also that a larger number of reconstruction modes can also be used to eliminate aliasing from the measurement with non-regular arrays, as described in [10] and referred to as under-modeling [11]. The idea consists in reconstructing the aberrated wavefront using more than the M terms of interest and then discarding the highest-order supposedly unnecessary coefficients, which are responsible for the main aliasing error included in the first M modes.

It is important to emphasize that the statistics used in this work is not complete to generate statistically robust ocular aberrations, once it was derived directly from the final results presented by Porter et al. [9] in which the measured Zernike coefficients for each member of the analyzed population are not available. Because of this limitation, one cannot generalize the obtained results. However, as more complete statistical models are available, the whole procedure can be repeated and a new optimized array can be specified. Therefore, the used statistics is to be understood as an example of application and as a first approximation for ocular aberrations.

The genetic algorithm optimization method (GA) [12] was chosen because it suits well the non-analytical H–S objective function, which exhibits a multitude of local minima. GA is a method to solve both constrained and unconstrained optimization problems and is based on natural selection, the process that drives biological evolution. It was implemented with MatLab© (R2008a), with the toolbox OPTIMTOOL—Optimization Toolbox 4.0. GA creates an initial set of 700 different arrays with the coordinates of microlenses randomly distributed, where the restrictions are that each lens should be within the array perimeter and that microlenses must not overlap. From this set, the 250 best arrays, in terms of lowest reconstruction error, compose the initial generation. Starting from the initial generation, the GA successively modifies the arrays from the current generation, referred to as parents, to create a new one, with the same number of arrays, which are called children, thus evolving towards an optimal solution. The new generation is always composed of three classes of children: elite, crossover and mutation. The elite children are passed from the previous generation without any changes. They correspond to the arrays which produced the two lowest reconstruction errors in the present generation. The other arrays are generated either through crossover or mutation operations. The crossover children are generated by a combination of a pair of parents, i.e. a combination of the microlens-coordinates of two arrays, whereas the mutation children are the result of small random changes in the parents. 70% of the arrays are generated through crossover and the others, through mutation and elitism. The specific algorithms to operate crossover and mutation can be selected in the MatLab© tool. In the present work, the used crossover function was *intermediate*, which creates children by a random weighted average of the parents. The mutation function was the *adaptive feasible*, which starts with a large mutation possibility in the beginning of the optimization, and means larger changes in the microlenses positions, and then sets it smaller as the process converges to the final solution.

The steps to find the optimized microlens array are summarized in the following outline:

i. Five different aberrated wavefronts were generated based on the statistics by Porter et al. [9]: four aberrations randomly produced by WG and one aberration featuring mean-valued Zernike coefficients. For a given microlens-array geometry, the H–S model computes the reconstruction error for each of these five aberrations, from which an average error is obtained. The average error is the parameter to be minimized by the GA.
ii. The GA creates the first generation of arrays with random distributions of microlenses. For each array, it computes the average error over the five aberrations mentioned earlier. The error values are used by the algorithm to define the expectation with which each array can be selected as a parent, so that the array with the smallest error is more likely to be chosen. Based on these values, the GA selects arrays to be used in the creation of the next generation, through the operations of elite, mutation and crossover. This step is done successively, until the average variation in the error over the 10 last generations is smaller than $1.58\times10^{-11}\lambda$. Then, the final result consists of the array that generated the smallest average error, which corresponds to a minimum in the objective function.
iii. The previous step was executed 10 times. Each of them generated a microlens array associated with a local minimum of the objective function. Therefore, the final result consists of 10 different microlens arrays.
iv. Each of these arrays was subjected to a comparison test with an orthogonal 6×6 array. Both the orthogonal and the optimized arrays have a square perimeter with 6 mm side and share the same microlens characteristics. The test consisted in estimating and comparing the reconstruction error for an arbitrarily fixed set of 2000 different aberrated wavefronts within the chosen ocular statistics. Among the previous set of 10 optimized arrays, the one that generated the best result was selected.

Although it cannot be concluded that this is the best array that can ever be found for this statistics, it is better than the regular one and, as a matter of fact, it is also better than at least other nine non-regular arrays. In practice, however, the obtained array is of great importance, once it can be readily fabricated and used in the optical setup to reduce the total reconstruction error introduced by the system. Whenever a new scenario has to be studied, the optimized array can be specified for the new statistics that describe the typical optical aberrations of interest.

Diaz-Santana et al. [13] present analytical expressions to calculate the reconstruction error for a known statistics. This model cannot be used in the current paper because it requires the measured Zernike coefficients for each member of the statistics population in order also to evaluate the correlation between Zernike terms. This is not available in the statistical data by Porter et al. [9] used here. If the statistics for description of the aberrations is known in detail, the performance of the arrays can be assessed through the model proposed by Diaz-Santana et al. [13], rather than testing the performance using a fixed and finite set of individual wavefront aberrations.

Before running the optimization process, some parameters used to describe the whole system had to be set, such as number of Zernike terms, wavelength (λ), lateral size of the microlens array, number of microlenses, microlens diameter and contour, refraction index and radius of curvature.

The microlens array was chosen to feature a square perimeter with a lateral dimension of 6 mm to exactly circumscribe the typical open-pupil diameter in the human eye. The tested arrays consisted of 10, 16 and 36 circular, plano-convex microlenses. It is important to emphasize that, when 20 Zernike terms are used, the minimum possible number of microlenses in the array is 10. This limitation is imposed by the implementation of the least-squares method in the modal reconstructor, which requires the computation of the Moore–Penrose pseudoinverse of the $2N \times M$ reconstruction matrix **B** (Eq. (3)). The condition to calculate the pseudoinverse is that the number of Zernike modes (M) and the number of microlenses (N) are related as $2N \geq M$ [11]. To guarantee some degree of freedom for the lens-position randomization, the microlens diameter was set to 750 μm. The aberrated wavefronts are sampled over the whole microlens surface and therefore the calculated local tilts correspond to the spatial average of the aberration over each microlens area, according to the lens model previously presented.

The used wavelength was 633 nm. The microlens refraction index at this wavelength was set to 1.4, to approximate that of commercially available materials for microlens fabrication (e.g., PMMA). The focal length of 4 cm corresponds to a radius of curvature of 16,000 μm, which, with a lens diameter of 750 μm, is close to the minimum achievable with the microlens-array fabrication capabilities of our photolithographic equipment and etching process [14,15]. These values have been set so to later guarantee the possibility of fabrication of the optimized microlens array.

3. Results and discussion

The optimization methodology described was applied to three cases: arrays with 10, 16 and 36 microlenses. The best arrays found in all cases were compared to the 16, 25 and 36-microlenses orthogonal arrays. The average values and the standard deviations for each case, calculated over 2000 aberrated wavefronts, are shown in Table 1. Fig. 3 illustrates the performance of the optimized 10-, 16- and 36-microlens arrays and the orthogonal 16-, 25- and 36-microlens arrays over 2000 aberrated wavefronts.

The results show that an optimized array can afford fewer microlenses, maintaining a low reconstruction error. As observed before, the reconstruction error comprises both sampling and numerical errors. Strikingly, through this optimization procedure both the reconstruction error and the number of lenses can be concomitantly reduced. The reduction in the reconstruction error though optimization can be clearly noted in the 16-microlens array. The same order of reduction could not be observed in the optimization of the 36-microlens array, once the orthogonal array is already capable of reconstructing the aberrated wavefront with a very low sampling error. Yoon [16] presents a graph that indicates that the maximum number of Zernike terms, which can be reliably reconstructed, is approximately the same as the number of microlenses in an orthogonal array. That means that if 20 Zernike terms need to be calculated, as it is the case in the present work, at least 20 microlenses are required for a reliable reconstruction. Using more than 20 microlenses may decrease the sampling error and, therefore, the reconstruction error, at least until the numerical errors become of the same order as the sampling error. In this sense, an orthogonal 16-microlens array cannot reliably reconstruct a wavefront aberration described by 20 Zernike coefficients. The optimization results demonstrate that it is possible to specify the distribution pattern of the 16 microlenses in the array so that it can generate reconstruction errors as low as the ones generated by the 25- and 36-microlens orthogonal arrays and moreover that arrays with even fewer microlenses, such as 10, can also be used to generate small reconstruction errors as the 25- and 36-microlens orthogonal arrays.

Table 1
Reconstruction errors generated by optimized and orthogonal arrays over 2000 wavefront aberrations.

#Lenslets (array geometry)	Average RMS reconstruction error	Standard deviation
10 (optimized)	$1.32 \times 10^{-8}\lambda$	$2.84 \times 10^{-8}\lambda$
16 (optimized)	$1.33 \times 10^{-8}\lambda$	$2.85 \times 10^{-8}\lambda$
36 (optimized)	$1.36 \times 10^{-8}\lambda$	$2.90 \times 10^{-8}\lambda$
16 (orthogonal)	$7.37 \times 10^{-2}\lambda$	$3.80 \times 10^{-2}\lambda$
25 (orthogonal)	$1.36 \times 10^{-8}\lambda$	$2.91 \times 10^{-8}\lambda$
36 (orthogonal)	$1.37 \times 10^{-8}\lambda$	$2.93 \times 10^{-8}\lambda$

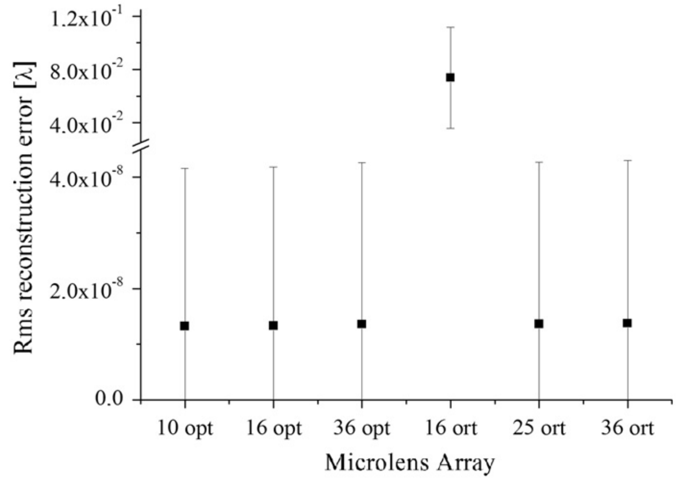

Fig. 3. Comparison between 10-, 16- and 36-microlens optimized (opt) arrays and 16-, 25- and 36-microlens orthogonal (ort) arrays. The used wavelength was λ=633 nm.

An array with fewer microlenses yields less data from the wavefront sensor, reducing processing time and round-off errors. Also, for a fixed array area, the microlens diameter can be maximized, improving the signal-to-noise ratio in the photodetection step and reducing the sizes of the resulting spots at the lenses foci. Using the fabrication process proposed by D.W. de Lima Monteiro et al. [14], even for non-regular lens centers, a 100% array fill factor can be maintained. Nevertheless, any modification of lens diameter and perimeter can alter the array performance and needs to be carefully assessed. The size of the microlens does not affect the sampled spatial-frequency spectrum, but influences the modulation transfer function (MTF) of the array, i.e. how accurately the amplitude of each frequency component is reproduced.

It is also important to compare the arrays considering their impact on the numerical calculation of the pseudoinverse of the reconstruction matrix **B** (Eq. (3)). This matrix is directly affected by the microlens positions in the array, since it contains the first derivatives of the Zernike functions evaluated at the respective sampling regions. The pseudoinverse calculation requires the inversion of the square matrix $\mathbf{B}^T\mathbf{B}$, which should be well conditioned. As pointed out by Navarro et al. [17], the 2-norm condition number can be computed to check the numerical stability of the inversion of a matrix. If the inverse of the condition number is used, it is always between 0 and 1. A condition number close to 1 means the matrix is well-conditioned. Moreover, Navarro et al. [17] states that the rank of the matrix to be inverted must equal the number of Zernike terms M to guarantee the matrix is not singular, i.e. $Rank(\mathbf{B}^T\mathbf{B})=20$. In this work, the condition number and the rank of the matrix $\mathbf{B}^T\mathbf{B}$ were calculated for the 10-, 16- and 36-microlens optimized arrays and

Table 2
Condition number and rank of the matrix $\mathbf{B}^T\mathbf{B}$.

#Lenslets (array geometry)	Condition number	Rank
10 (optimized)	8×10^{-4}	20
16 (optimized)	2×10^{-3}	20
36 (optimized)	7×10^{-3}	20
16 (orthogonal)	2×10^{-17}	18
36 (orthogonal)	1×10^{-2}	20

for the 16- and 36-microlens orthogonal arrays. The results are shown in Table 2.

The rank of the 16-microlens orthogonal array was 18, meaning that matrix $\mathbf{B}^T\mathbf{B}$ is singular, or close to singular, for this case and indicates this array is not appropriate to sample Zernike polynomials. On the other hand, the rank of the other arrays, including the 16-microlens optimized array, was 20, which equals the total number of used Zernike terms. This guarantees that the matrix is not singular when these arrays are used and therefore the matrix $\mathbf{B}^T\mathbf{B}$ can be inverted. The values of condition numbers of all the arrays, except for the 16-microlens orthogonal one, indicate the respective matrices are well-conditioned. That guarantees the optimized array preserves accuracy and numerical stability in the wavefront reconstruction [17].

4. Conclusions

This work proposed a methodology to numerically optimize the microlens-array geometry for a Hartmann–Shack (H–S) wavefront sensor, aiming at specifying arrays with minimum number of microlenses whilst maintaining reconsctruction errors as low as the ones generated by arrays with more sampling microlenses. The method consists in using the genetic algorithm optimization approach to minimize the reconstruction error of the H–S wavefront sensor by changing the coordinates of microlenses in the array. The wavefronts used to calculate the reconstruction error need to be obtained from a statistical description of the distortions in a particular application. In particular, the method was applied to the ophthalmological context, as an example of application, it was shown that an optimized 10-microlens array generated smaller reconstruction errors than an orthogonal 36-microlens array.

With slight modifications to the H–S model this method can also be used to optimize a Hartmann-mask pattern or the probe-laser coordinates in laser ray tracing [4].

This methodology is generic and can be applied to other wavefront statistics as well. An optimized microlens array can be found to either comply with the maximum tolerated error in that particular context with a reduced number of microlenses, or to yield a lower error altogether.

Acknowledgement

This work has been supported by the Brazilian National Council for Technological and Scientific Development (CNPq) and Research Support Foundation of the State of Minas Gerais (FAPEMIG).

References

[1] Platt BC, Shack R. History and principles of Shack–Hartmann wavefront sensing. J Refract Surg 2001;17:573–7.
[2] Greenaway A, Burnett J. Industrial and medical applications of adaptive optics. IOP 2004.
[3] Llorente L, Marcos S, Dorronsoro C, Burns SA. Effect of sampling on real ocular aberrations measurements. J Opt Soc Am A 2007;9:2783–96.
[4] de Lima Monteiro DW. CMOS-based integrated wavefront sensor. Delft University Press; 2002.
[5] Born M, Wolf E. Principles of optics: electromagnetic theory of propagation, interference, and diffraction of light. Pergamon Press; 1989.
[6] Soloviev O, Vdovin G. Hartmann–Shack test with random masks for modal wavefront reconstruction. Opt Express 2005;13:9570–84.
[7] Noll RJ. Zernike polynomials and atmospheric turbulence. J Opt Soc Am 1976;66:207–11.
[8] Cubalchini R. Modal wave-front estimation from phase derivative measurements. J Opt Soc Am 1979;69:972–7.
[9] Porter J, Guirao A, Cox IG, Williams DR. Monochromatic aberrations of the human eye in a large population. J Opt Soc Am 2001;18:1793–803.
[10] de Lima Monteiro DW, Furtado ASO, Soloviev O, Ferreira Jr. AI, Teixeira FB, Melo JGM, Vdovin G. Improvement of dynamics features of an adaptive optical system for detection and correction of ophthalmic distortions. In: NSTI-Nanotech 3, 2006. p. 413–8.
[11] Wang JY, Silva DE. Wave-front interpretation with Zernike polynomials. Appl Opt 1980;19:1510–8.
[12] Mitchell M. An introduction to genetic algorithms. The MIT Press; 1999.
[13] Diaz-Santana L, Walker G, Bará SX. Sampling geometries for ocular aberrometery: a model for evaluation of performance. Opt Express 2005;13:8801–18.
[14] de Lima Monteiro DW, Akhzar-Mehr O, Sarro PM, Vdovin G. Single-mask microfabrication of aspherical optics using KOH anisotropic etching of Si. Opt Express 2003;18:2244–2252.
[15] de Lima Monteiro DW, Akhzar-Mehr O, Vdovin G. Prime microlens arrays for Hartmann–Shack sensors—an economical fabrication technology. In: 4th adaptive optics conference for industry and medicine. Springer; 2003. p. 197–205.
[16] Yoon G. Wavefront sensing and diagnostic uses. In: Porter J, Queener HM, Lin JE, Thorn K, Awwal A, editors. Adaptive optics for vision science: principles, practices, design and applications. John Wiley & Sons; 2006. p. 63–81.
[17] Navarro R, Arines J, Rivera R. Direct and inverse discrete Zernike transform. Opt Express 2009;17:24269–81.